REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
JULY 2010	Conf Paper (Briefing Charts)	April 2008 – January 2009

4. TITLE AND SUBTITLE

SEMANTIC SEARCH WITH SELF-MAINTAINING CLASSES

5a. CONTRACT NUMBER
FA8750-08-C-0157

5b. GRANT NUMBER
N/A

5c. PROGRAM ELEMENT NUMBER
65502F

6. AUTHOR(S)

Terry Patten and Stephen Hookway

5d. PROJECT NUMBER
08SB

5e. TASK NUMBER
IR

5f. WORK UNIT NUMBER
75

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Charles River Analytics, Inc.
625 Mount Auburn Street, Suite 3
Cambridge, MA 02138-4555

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFRL/RIEB
525 Brooks Road
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)
N/A

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-RI-RS-TP-2010-33

12. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED. PA #: 88ABW-2010-3417
Date Cleared: 21-June-2010

13. SUPPLEMENTARY NOTES
This work was funded in whole or in part by Department of the Air Force contract number FA8750-08-C-0157. The U.S. Government has for itself and others acting on its behalf an unlimited, paid-up, nonexclusive, irrevocable worldwide license to use, modify, reproduce, release, perform, display, or disclose the work by or on behalf of the Government.

14. ABSTRACT
The SBIR contract entitled Knowledge-based Access and Data Integration (KADI) resulted in a paper being presented at the 2010 International Conference on Semantic Web and Web Services held at Las Vegas, NV on 12-15 July-2010.

15. SUBJECT TERMS
Semantic Integration, Imagery, Intelligence Surveillance Reconnaissance, Service Oriented Architecture

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Fred N. Haritatos
U	U	U	UU	12	19b. TELEPHONE NUMBER (*Include area code*) N/A

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Semantic Search With Self-Maintaining Classes

Prepared for:

SWWS 2010
Las Vegas

Presented by:

Terry Patten

Stephen Hookway

This work was supported by US Air Force Contract FA8750-08-C-0157

12-15 July 2010

www.cra.com

Background

- Semantic search is all about *inferring* membership in query classes
- Traditionally, semantic search has focused on class-subclass inference (subsumption)
  - Query: Aircraft
  - Match: MiG-23
    - Subclass of Fighter
      - Subclass of MilitaryAircraft
        - Subclass of Aircraft
- Most often, classes are enumerated through subsumption trees
- But class membership can also be inferred using
  - Logic (a class is the intersection of several other classes)
  - Properties (a class includes any *x* that has property *y*)
- In general, class membership descriptions can be complex
  - In OWL 2 these are called "Class Expressions"
  - In Description Logic these are called "Complex Concepts"

Class Expressions and Semantic Search

- Inferring class membership using class expressions is a powerful semantic search technique because:
  - It provides a mechanism loosely analogous to database "views"
    - Alternate ways of organizing information can be built on top of an existing ontology
  - Query classes do not have to reflect the subsumption hierarchy
  - Query classes do not have to be enumerated
- The inferred classes are self-maintaining because their membership automatically changes as the membership of underlying classes changes

A Case Study: Threats to Air Operations

- Problem: Air Force analysts searching image metadata for "threats to air operations"
- Currently, these searches are done using keywords
- But there may be many names for weapon systems:
  - NATO designations
  - Coalition partner names
  - Anglicized manufacturer names
- Having to include all these names in queries places a burden on the analyst
  - An ontology can keep track of all these names
  - This is an old technique: using a thesaurus to expand a query
- But even with only one term for each weapon, there are still a lot of weapons to list
- We would like to search for whole classes of threats

Searching for Threat Classes

- Threats to air operations include members of traditional weapons classes such as:
  - Aircraft
  - Missiles
  - Artillery
- But not all members of these traditional military classes are threats to air operations
  - Bombers are aircraft, but are not really a threat to an air operation
  - Ballistic missiles are not a threat
  - Surface-to-surface artillery is not a threat
- It is tempting to enumerate the threats to air operations
  - What are all the types of Threats from the air …
  - What are all the types of Threats from the surface …
- A better idea is to *infer* threats based on weapon properties

Inferring Threats from Properties

- Using OWL-DL we can say that the class of weapons that fires air-to-air missiles is a subclass of threats to air operations
- We don't have to say (explicitly) what weapons are in that class—they are inferred based on the property
- We don't have to know (or care) where in the ontology such weapons exist
- This technique allows us to cut across existing subsumption hierarchies and create alternate views of the knowledge
- In particular, we can create views that line up well with useful queries

Example: BM-21

- The BM-21 is a multiple-rocket launcher
  - Classic surface-to-surface artillery
- But adversaries modified the rockets
  - Used timed-fuzes
- And pointed the weapon straight up
  - Rockets become heavy flak


BM-21 in Afghanistan

- The list of munitions that the BM-21 fires is updated to include these modified, now anti-aircraft rockets
- Since weapons that fire anti-aircraft artillery munitions are a subclass of threats to air operations…
  - The BM-21 *automatically* becomes a threat as soon as that property appears

OWL-DL Class Expression that Matches BM-21

```
<owl:Class rdf:about="#FunctionalAntiAircraftArtillery">
      <rdfs:subClassOf rdf:resource="#ThreatToAirOperations"/>

      <rdfs:subClassOf rdf:resource="#Artillery"/>
      <owl:equivalentClass>
            <owl:Restriction>
               <owl:onProperty rdf:resource="#Fires"/>
               <owl:someValuesFrom rdf:resource="#AntiAircraftMunitions"/>
            </owl:Restriction>
       </owl:equivalentClass>
</owl:Class>
```

- This class expression defines—and *maintains*—a useful class
  - Prior to having anti-aircraft munitions, the BM-21 it is not a member of this class
  - Once the BM-21 fires anti-aircraft munitions it *automatically* becomes a member of this class

Case Study Summary

- Weapons are threats to air operations if they
  - Fire air-to-air missiles
  - Fire surface-to-air missiles
  - Fire anti-aircraft artillery munitions
  - Etc.
- We define a class expression for each of these
- And make them all subclasses of "ThreatsToAirOperations"
- Weapons automatically become members of the Threat class if they match the class expressions
- The class expressions automatically maintain class membership
- For scalable OWL-DL, we are implementing our semantic search using PelletDb (http://clarkparsia.com/pelletdb/)
  - A tight integration of Pellet and Oracle Semantic Technology

Generality of Search Using Self-Maintaining Classes

- This technique can be used in *any* domain where some query classes do not line up well with the subsumption hierarchy
- Some non-military possibilities:
  - Find me all the risky investments …
  - Find me all the high-performance cars …
  - Find me all the dangerous substances …
  - …

Questions?

charles river analytics

+ Charles River Analytics Inc.
- 625 Mount Auburn St.
- Cambridge, MA 02138
- p: 617.491.3474
- f: 617.868.0780
- www.cra.com